

A RATE OF CONVERGENCE FOR THE ARCSINE LAW BY STEIN'S METHOD

CHRISTIAN DÖBLER

ABSTRACT. Using Stein's method for the Beta distributions and a recent technique by Goldstein and Reinert of comparing the Stein characterization of the target distribution with that of the approximating distribution we prove a rate of convergence in the classical arcsine law, which states that the distribution of the relative time spent positive by a symmetric random walk on \mathbb{Z} converges weakly to the arcsine distribution on $[0, 1]$.

1. INTRODUCTION

Consider a game between two players A and B that consists of consecutive tossings of a fair coin. Each time the coin shows heads, player A has to pay one dollar to player B and conversely, each time the coin shows tails, player A obtains one dollar from player B . If we consider the process that gives for each discrete time n the current fortune of player A , then by the symmetry of the model one is led to the conjecture, that for n sufficiently large, the relative amount of time, that player A is in the lead should be roughly one-half. The so-called (first) arcsine law states, that this intuition is entirely wrong. In fact, it is more likely that one of the players will lead for nearly all of the time.

Let $(S_k)_{k \geq 0}$ be the symmetric random walk on \mathbb{Z} , i.e. we have $S_k := \sum_{j=1}^k \varepsilon_j$ ($k \geq 0$) for i.i.d. random variables $\varepsilon_1, \varepsilon_2, \dots$ with $P(\varepsilon_1 = 1) = P(\varepsilon_1 = -1) = \frac{1}{2}$. Letting $X_j := 1_{\{S_{j-1} \geq 0, S_j \geq 0\}}$, $T_m := \sum_{j=1}^{2m} X_j$, $R_m := \frac{1}{2}T_m$ and $W_m := \frac{1}{2m}T_m = \frac{1}{m}R_m$ it is a classical result, first proven by Paul Lévy for Brownian motion, that as $m \rightarrow \infty$ we have

$$\mathcal{L}(W_m) \xrightarrow{\mathcal{D}} \nu_{\frac{1}{2}, \frac{1}{2}},$$

where, for $a, b > 0$ $\nu_{a,b}$ denotes the *Beta distribution* to the parameters a and b on $[0, 1]$, which has density $q_{a,b}(x) := \frac{1}{B(a,b)} x^{a-1} (1-x)^{b-1} 1_{(0,1)}(x)$ with $B(a,b)$ denoting

The author has been supported by Deutsche Forschungsgemeinschaft via SFB-TR 12. Ruhr-Universität Bochum, Fakultät für Mathematik, NA 3/68, D-44780 Bochum, Germany. christian.doebler@ruhr-uni-bochum.de

Keywords: Stein's method, arcsine law, symmetric random walk, Chung-Feller Theorem .

the *Beta function*. Consequently, $\nu := \nu_{\frac{1}{2}, \frac{1}{2}}$ is the arcsine distribution on $[0, 1]$ with density $q(x) := q_{\frac{1}{2}, \frac{1}{2}}(x) = \frac{1}{\pi} \frac{1}{\sqrt{x(1-x)}} 1_{(0,1)}(x)$. A proof of this well-known theorem can be found for example in [Fel68].

Recently, there has been some progress in Stein's method for the family of Beta distributions by Goldstein and Reinert (see [GR12]) and by the author of the present article (see [Döb12]). Both of these preprints prove rates of convergence in a Polya urn model. In this paper we will use the general results from the preprints [GR12] and [Döb12] and especially the technique of comparing the Stein characterization of the target distribution with that of the (discrete) approximating distribution from [GR12] to prove bounds on the Wasserstein distance of the distribution $\mathcal{L}(W_m)$ of W_m to the arcsine distribution ν .

2. STEIN'S METHOD FOR THE ARCSINE DISTRIBUTION AND FOR $\mathcal{L}(R_m)$

Stein's method is very useful tool for proving distributional convergence. Its main advantage over other techniques is, that it automatically yields concrete error bounds on various distributional distances. Since its introduction in 1972 in the seminal paper [Ste72] by Charles Stein for the univariate standard normal distribution, there has been much progress in adapting Stein's idea of linking a characterizing operator for the target distribution to a differential equation (in the absolutely continuous case) or to a difference equation (in the discrete case), the *Stein equation*, to other distributions, as for example the Poisson distribution (see [Che75]), the Gamma distribution (see [Luk94]), the exponential distribution (see [CFR11] and [PR11]), the geometric distribution (see [PRR]) and many others. For a general introduction to Stein's method we refer to the book [CGS11] which emphasizes normal approximation but also treats approximation by other distributions. Here, we will make use of the recent development of Stein's method for the Beta distributions (specialized to the arcsine distribution), which was done independently and with different emphases in [GR12] and [Döb12]. Furthermore, we will heavily use the approach from [GR12] for finding various Stein characterizations of a distribution supported on \mathbb{Z} . This topic was also explored in [LS].

We start with Stein's method for the arcsine distribution ν . The following result, a slight variant of Proposition 3.1 in [Döb12], gives a Stein characterization for ν . We denote by \mathcal{K} the class of all continuous and piecewise continuously differentiable functions $f : \mathbb{R} \rightarrow \mathbb{R}$ vanishing at infinity with $\int_{\mathbb{R}} |f'(x)| \sqrt{x(1-x)} dx < \infty$.

PROPOSITION 2.1. *A real-valued random variable X has the arcsine distribution ν on $[0, 1]$ if and only if for all functions $f \in \mathcal{K}$ the expected values $E[X(1-X)f(X)]$ and $E[(X - \frac{1}{2})f(X)]$ exist and coincide.*

According to Stein's idea, by Proposition 2.1, for a given ν -integrable test function $h : \mathbb{R} \rightarrow \mathbb{R}$ one is led to consider the Stein equation

$$(1) \quad x(1-x)f'(x) + \left(\frac{1}{2} - x\right)f(x) = h(x) - \nu(h),$$

which is to be solved for the unknown function f of $x \in \mathbb{R}$ (or, at least, $x \in [0, 1]$). From the theory in Section 3 of [Döb12] we know that there exists a unique solution f_h to (1), defined on \mathbb{R} , which is bounded on $[0, 1]$. For $x \in (0, 1)$ it is given by

$$(2) \quad f_h(x) = \frac{1}{x(1-x)q(x)} \int_0^x (h(t) - \nu(h))q(t)dt = \frac{-1}{x(1-x)q(x)} \int_x^1 (h(t) - \nu(h))q(t)dt.$$

Furthermore, f_h is continuous at 0 and 1 as long as h is. The following result is a special case of Proposition 3.7 in [Döb12].

LEMMA 2.2. *Let $h : \mathbb{R} \rightarrow \mathbb{R}$ be Borel-measurable and ν -integrable.*

- (a) *If h is bounded, then $\|f_h\|_\infty \leq 2\|h - \nu(h)\|_\infty$.*
- (b) *If h is Lipschitz, then $\|f_h\|_\infty \leq 2\|h'\|_\infty$ and $\|f'_h\|_\infty \leq C_1\|h\|_\infty$, where the finite constant C_1 does not depend on h .*
- (c) *If h is twice differentiable with bounded first and second derivative, then $\|f''_h\|_\infty \leq C_2(\|h'\|_\infty + \|h''\|_\infty)$, where the finite constant C_2 does not depend on h .*

Proof. Noting that $\frac{1}{2}$ is the median for ν with $q(1/2) = \frac{2}{\pi}$, this follows immediately from Proposition 3.7 in [Döb12] with $a = b = \frac{1}{2}$. \square

REMARK 2.3. Note that the bounds on the Stein solutions f_h from Lemmas 3.3-3.5 in [GR12] do not cover the case of the arcsine distribution, since they all impose the condition $a, b \in [1, \infty)$. It is the restriction to these parameters, that allows the authors to provide explicit constants in place of C_1 from Lemma 2.2 which furthermore yield explicit constants for the rate of convergence in the Polya urn model.

Now, we turn to a suitable version of Stein's method for the distribution $\mathcal{L}(R_m)$ of R_m . To do so, we will first give an explicit formula for the probability mass function $p(k)$ of R_m , which is given by a famous Theorem by Chung and Feller. First, however, we will need the following lemma.

LEMMA 2.4. *Let m be a positive integer. Then*

- (a) *$X_{2j-1} = X_{2j}$ for $j = 1, \dots, m$*
- (b) *T_m has values in $2 \cdot \{0, \dots, m\}$ and hence R_m has values in $\{0, \dots, m\}$.*

- (c) Letting $\tilde{X}_j := 1 - X_j$ we have $\tilde{X}_j = 1_{\{S_{j-1} \leq 0, S_j \leq 0\}}$ and $(\tilde{X}_1, \dots, \tilde{X}_n) \stackrel{\mathcal{D}}{=} (X_1, \dots, X_n)$.

Proof. We prove (a) by induction on j . It is easy to see, that $X_1 = X_2$ always holds. Now let $1 \leq j \leq m-1$. Then we have $X_{2j-1} = X_{2j}$ by the induction hypothesis. Suppose, that $X_{2j-1} = X_{2j} = 1$. If $S_{2j} = 0$, then the claim $X_{2j+1} = X_{2j+2}$ follows in the same manner as $X_1 = X_2$. If $S_{2j} > 0$, then necessarily $S_{2j} \geq 2$, yielding $S_{2j+1} > 0$ and $S_{2j+1} \geq 0$. Hence, $X_{2j+1} = X_{2j} = 1$ in this case. If, contrarily, $X_{2j-1} = X_{2j} = 0$, the proof is similar.

Assertion (b) follows immediately from (a). The first assertion from (c) is clear since either both, S_{2j-1} and S_{2j} , are nonnegative or nonpositive. Now, observe, that there is a (measurable) function f such that $(X_1, \dots, X_n) = f(S_1, \dots, S_n)$. Since $\tilde{X}_j = 1_{\{-S_{j-1} \geq 0, -S_j \geq 0\}}$ and $(S_1, \dots, S_n) \stackrel{\mathcal{D}}{=} (-S_1, \dots, -S_n)$ by symmetry, we have

$$(\tilde{X}_1, \dots, \tilde{X}_n) = f(-S_1, \dots, -S_n) \stackrel{\mathcal{D}}{=} f(S_1, \dots, S_n) = (X_1, \dots, X_n).$$

□

The proof of the following well-known theorem can be found for example in [Fel68].

THEOREM 2.5 (Chung-Feller Theorem). *Let m be a positive integer. Then, for each $0 \leq k \leq m$ we have*

$$P(R_m = k) = P(T_m = 2k) = u_{2k}u_{2m-2k},$$

where $u_0 := 1$ and $u_{2j} := \binom{2j}{j}2^{-j}$ for $j \geq 1$ denotes the probability that the symmetric random walk returns to zero at time $2j$.

Thus, by Theorem 2.5 the probability mass function $p : \mathbb{Z} \rightarrow \mathbb{R}$ corresponding to R_m is given by $p(k) = 0$ for $k \in \mathbb{Z} \setminus \{0, \dots, m\}$ and by

$$(3) \quad p(k) = u_{2k}u_{2m-2k} = 2^{-2m} \binom{2k}{k} \binom{2m-2k}{m-k}$$

for $0 \leq k \leq m$.

In the following, we review the recent adaption of the so-called *density approach* for absolutely continuous distributions (see, e.g. [SDHR04], [EL10], [CS11] and [CGS11]) to discrete distributions on the integers, which was done in [GR12] and also in [LS]. For reasons of simplicity we restrict ourselves to the case of finite integer intervals. A finite integer interval is a set I of the form $I = [a, b] \cap \mathbb{Z}$ for some integers $a \leq b$. Given a probability mass function $p : \mathbb{Z} \rightarrow \mathbb{R}$ with $p(k) > 0$ for $k \in I$ and $p(k) = 0$ for $k \in \mathbb{Z} \setminus I$, we consider the function $\psi : I \rightarrow \mathbb{R}$ given by the formula

$$(4) \quad \psi(k) := \frac{\Delta p(k)}{p(k)},$$

where for a function f on the integers $\Delta f(k) := f(k+1) - f(k)$ denotes the *forward difference operator*. Note that by definition always $\psi(b) = -1$ if $I = [a, b] \cap \mathbb{Z}$. For such a probability mass function p with support a finite integer interval $I = [a, b] \cap \mathbb{Z}$, let $\mathcal{F}(p)$ denote the class of all real-valued functions f on \mathbb{Z} such that $f(a-1) = 0$. The following result is a special case of Proposition 2.1 of [GR12].

PROPOSITION 2.6. *Let Z be a \mathbb{Z} -valued random variable with probability mass function p which is supported on the finite integer interval $I = [a, b] \cap \mathbb{Z}$ and is positive there. Then, a given random variable X with support I has the probability mass function p if and only if for all $f \in \mathcal{F}(p)$ it holds that*

$$(5) \quad E[\Delta f(X-1) + \psi(X)f(X)] = 0.$$

The next result, a version of Corollary 2.1 from [GR12], yields various other Stein characterizations for the distribution corresponding to p from Proposition 2.6.

COROLLARY 2.7. *Let Z be a \mathbb{Z} -valued random variable with probability mass function p which is supported on the finite integer interval $I = [a, b] \cap \mathbb{Z}$ and is positive there. Let $c : [a-1, b] \cap \mathbb{Z} \rightarrow \mathbb{R} \setminus \{0\}$ be an arbitrary function. Then, in order that a given random variable X with support I has the probability mass function p it is necessary and sufficient that for all functions $f \in \mathcal{F}(p)$ we have*

$$(6) \quad E[c(X-1)\Delta f(X-1) + [c(X)\psi(X) + \Delta c(X-1)]f(x)] = 0.$$

REMARK 2.8. Letting $\gamma(k) := c(k)\psi(k) + \Delta c(k-1)$ we see that c satisfies the difference equation $\Delta c(k-1) = \gamma(k) - c(k)\psi(k)$. This exactly corresponds to the differential equation $\eta'(x) = \gamma(x) - \eta(x)\psi(x)$ from Formula (14) in [Döb12], where $\psi(x) := \frac{p'(x)}{p(x)}$ is the logarithmic derivative of the density p . In [Döb12] it is shown, that this differential equation must hold, in order that a given distribution μ with density p satisfies the Stein identity $E[\eta(Z)g'(Z) + \gamma(Z)f(Z)] = 0$, where $Z \sim \mu$. So also in this respect, there is a strong analogy between the absolutely continuous and the discrete case.

Now, with the abstract results at hand, we return to the concrete distribution of R_m which has probability mass function p supported on $I := [0, m] \cap \mathbb{Z}$ and given by (3). Using the relation

$$\binom{2j}{j} = \frac{4j-2}{j} \binom{2j-2}{j-1} \quad \text{for } j \geq 1$$

it can easily be checked, that in this case ψ is given by

$$(7) \quad \psi(k) = \frac{2k - m + 1}{(k+1)(2(m-k) - 1)}, \quad 0 \leq k \leq m.$$

This motivates the definition $c(k) := (k+1)(2(m-k) - 1)$ for $k = 0, \dots, m$. These observations lead to the following lemma.

LEMMA 2.9. *Let p be the probability mass function of R_m as given by (3). A random variable X with support $I = [0, m] \cap \mathbb{Z}$ has probability mass function p if and only if for all $f \in \mathcal{F}(p)$ it holds that*

$$(8) \quad E \left[X \left((m - X) + \frac{1}{2} \right) \Delta f(X - 1) + \left(\frac{m}{2} - X \right) f(X) \right] = 0.$$

Proof. This follows from Corollary 2.7, since by the definition of c we have

$$c(k)\psi(k) + \Delta c(k-1) = 2 \left(\frac{m}{2} - k \right)$$

and

$$c(k-1) = k(2(m-k) + 1) = 2k \left(m - k + \frac{1}{2} \right).$$

□

3. A RATE OF CONVERGENCE FOR THE ARCSINE LAW

In this section we will use the tools from Section 2 to prove a rate of convergence in the Wasserstein distance for the arcsine law. Recall, that for two distributions μ_1 and μ_2 on $(\mathbb{R}, \mathcal{B})$, whose first moments exist, the Wasserstein distance is given by

$$d_{\mathcal{W}}(\mu_1, \mu_2) := \sup_{h \in \text{Lip}(1)} \left| \int_{\mathbb{R}} h d\mu_1 - \int_{\mathbb{R}} h d\mu_2 \right|$$

and for two real-valued random variables X and Y one defines

$$d_{\mathcal{W}}(X, Y) := d_{\mathcal{W}}(\mathcal{L}(X), \mathcal{L}(Y)) = \sup_{h \in \text{Lip}(1)} |E[h(X)] - E[h(Y)]|,$$

where $\text{Lip}(1)$ denotes the class of all Lipschitz-continuous functions h on \mathbb{R} with minimal Lipschitz constant $\|h'\|_{\infty} \leq 1$. It is known, that on the space of probability

measures with existing first moment, convergence in the Wasserstein distance is stronger than weak convergence.

THEOREM 3.1. *There exists a finite constant $C > 0$ such that for each positive integer m we have*

$$d_W(\mathcal{L}(W_m), \nu) \leq \frac{C}{m}.$$

REMARK 3.2. To the best of my knowledge this is the first result that gives a rate of convergence of order m^{-1} for the arcsine law. The restriction to even times $n = 2m$ is immaterial and only for convenience, since the formula from the Chung-Feller Theorem only holds for these times. Since for odd times $n = 2m + 1$

$$\left| \frac{1}{2m+1} \sum_{j=1}^{2m+1} X_j - W_m \right| \leq \frac{2}{2m+1},$$

the same rate of convergence also holds for the whole sequence of positive times of the random walk.

It may be seen from the proof of Theorem 3.1, that the constant C can be made explicit in terms of the constant C_1 from Lemma 2.2.

Proof of Theorem 3.1. The proof follows the lines of the proof of Theorem 3.1 from [GR12] and is included for reasons of completeness. Using the notation from [GR12], for a function f and $y > 0$ let

$$\Delta_y f(x) := f(x+y) - f(x).$$

We also write $W := W_m$ and $R := R_m$. Let $h \in \text{Lip}(1)$ be fixed and let $f := f_h$ be the corresponding solution to the Stein equation (1) given by (2) for $x \in (0, 1)$ but which we set equal to zero for $x \in \mathbb{R} \setminus [0, 1]$. Consider the function $g(x) := f(x/m)$ which is zero on $\mathbb{R} \setminus [0, m]$. Then by Lemma 2.9 and upon dividing by m in (8) we obtain

$$\begin{aligned} 0 &= \frac{1}{m} E \left[R \left((m-R) + \frac{1}{2} \right) \Delta g(R-1) + \left(\frac{m}{2} - R \right) g(R) \right] \\ &= E \left[mW \left((1-W) + \frac{1}{2m} \right) \Delta_{1/m} f \left(W - \frac{1}{m} \right) + \left(\frac{1}{2} - W \right) f(W) \right] \end{aligned}$$

Inserting this into the Stein identity resulting from the stein equation (1) we obtain

$$\begin{aligned}
& E[h(W)] - \nu(h) \\
&= E\left[W(1-W)f'(W) + \left(\frac{1}{2} - W\right)f(W)\right] \\
&= E\left[W(1-W)f'(W) - mW\left((1-W) + \frac{1}{2m}\right)\Delta_{1/m}f\left(W - \frac{1}{m}\right)\right] \\
(9) \quad &= E\left[W(1-W)f'(W) - mW(1-W)\Delta_{1/m}f\left(W - \frac{1}{m}\right)\right] + E_1,
\end{aligned}$$

with

$$(10) \quad |E_1| = \frac{1}{2} \left| E\left[W\Delta_{1/m}f\left(W - \frac{1}{m}\right)\right] \right| \leq \frac{1}{2m} \|f'\|_\infty E[W] = \frac{1}{4m} \|f'\|_\infty \leq \frac{C_1}{4m}$$

by Lemma 2.2 (b) and by $E[W] = 1/2$, which follows for example from Lemma 2.4 (c) by symmetry. Using the fundamental theorem of calculus, we rewrite the remaining term in (9) as

$$\begin{aligned}
& E\left[W(1-W)f'(W) - mW(1-W)\Delta_{1/m}f\left(W - \frac{1}{m}\right)\right] \\
&= E\left[W(1-W)\left(f'(W) - m\int_{W-\frac{1}{m}}^W f'(t)dt\right)\right] \\
&= mE\left[\int_{W-\frac{1}{m}}^W W(1-W)\left(f'(W) - f'(t)\right)dt\right] \\
(11) \quad &= mE\left[\int_{W-\frac{1}{m}}^W \left(W(1-W)f'(W) - t(1-t)f'(t)\right)dt\right] + E_2
\end{aligned}$$

where

$$E_2 := -mE\left[\int_{W-\frac{1}{m}}^W \left(W(1-W) - t(1-t)\right)f'(t)dt\right]$$

and hence, again by the fundamental theorem of calculus,

$$\begin{aligned}
|E_2| &= m \left| E \left[\int_{W-\frac{1}{m}}^W f'(t) \int_t^W (1-2s) ds dt \right] \right| \\
&\leq m \left(1 + \frac{2}{m} \right) \|f'\|_\infty E \left[\int_{W-\frac{1}{m}}^W (W-t) dt \right] \\
&= (m+2) \|f'\|_\infty \int_0^{\frac{1}{m}} u du \\
(12) \quad &= \frac{m+2}{2m^2} \|f'\|_\infty \leq C_1 \frac{m+2}{2m^2}
\end{aligned}$$

where we have used Lemma 2.2 for the last step and the inequality $|1-2s| \leq (1+\frac{2}{m})$ for relevant values of s for the first step.

It remains to deal with the first expectation in (11). Since $f = f_h$ solves the Stein equation (1) and by the fundamental theorem of calculus for Lebesgue integration, we obtain

$$\begin{aligned}
(13) \quad &mE \left[\int_{W-\frac{1}{m}}^W \left(W(1-W)f'(W) - t(1-t)f'(t) \right) dt \right] \\
&= mE \left[\int_{W-\frac{1}{m}}^W \left(h(W) - \nu(h) - \left(\frac{1}{2} - W \right) f(W) - h(t) + \nu(h) + \left(\frac{1}{2} - t \right) f(t) \right) dt \right] \\
&= mE \left[\int_{W-\frac{1}{m}}^W \left(h(W) - h(t) + \left(W - \frac{1}{2} \right) f(W) - \left(t - \frac{1}{2} \right) f(t) \right) dt \right] \\
&= mE \left[\int_{W-\frac{1}{m}}^W \left(\int_t^W h'(s) ds + \int_t^W \left(f(s) + \left(s - \frac{1}{2} \right) f'(s) \right) ds \right) dt \right]
\end{aligned}$$

The inner integrals from (13) are bounded separately. As to the first one,

$$(14) \quad \left| mE \left[\int_{W-\frac{1}{m}}^W \int_t^W h'(s) ds dt \right] \right| \leq m \|h'\|_\infty E \left[\int_{W-\frac{1}{m}}^W (W-t) dt \right] = \frac{1}{2m} \|h'\|_\infty \leq \frac{1}{2m}.$$

For the second one, since $|\frac{1}{2} - s| \leq \frac{1}{m} + \frac{1}{2} \leq 3/2$ for the relevant values of s , we have

$$\begin{aligned}
& m \left| E \left[\int_{W-\frac{1}{m}}^W \int_t^W \left(f(s) + \left(s - \frac{1}{2} \right) f'(s) \right) ds dt \right] \right| \\
& \leq m \left(\|f\|_\infty + \frac{3}{2} \|f'\|_\infty \right) E \left[\int_{W-\frac{1}{m}}^W \int_t^W ds dt \right] \\
& = m \left(\|f\|_\infty + \frac{3}{2} \|f'\|_\infty \right) E \left[\int_{W-\frac{1}{m}}^W (W-t) dt \right] \\
& = m \left(\|f\|_\infty + \frac{3}{2} \|f'\|_\infty \right) \frac{1}{2m^2} \\
(15) \quad & \leq \left(2 + \frac{3}{2} C_1 \right) \|h'\|_\infty \frac{1}{2m} \leq \frac{4 + 3C_1}{4m},
\end{aligned}$$

where we have used Lemma 2.2 for the next to last inequality. Since $h \in \text{Lip}(1)$ was arbitrary, the conclusion of the theorem follows from (10), (12), (14) and (15). \square

REFERENCES

- [CFR11] Sourav Chatterjee, Jason Fulman, and Adrian Röllin. Exponential approximation by Stein's method and spectral graph theory. *ALEA Lat. Am. J. Probab. Math. Stat.*, 8:197–223, 2011.
- [CGS11] Louis H. Y. Chen, Larry Goldstein, and Qi-Man Shao. *Normal approximation by Stein's method*. Probability and its Applications (New York). Springer, Heidelberg, 2011.
- [Che75] Louis H. Y. Chen. Poisson approximation for dependent trials. *Ann. Probability*, 3(3):534–545, 1975.
- [CS11] Sourav Chatterjee and Qi-Man Shao. Nonnormal approximation by Stein's method of exchangeable pairs with application to the Curie-Weiss model. *Ann. Appl. Probab.*, 21(2):464–483, 2011.
- [Döb12] C. Döbler. Stein's method of exchangeable pairs for absolutely continuous, univariate distributions with applications to the Polya urn model. [arXiv:1207.0533](#), 2012.
- [EL10] Peter Eichelsbacher and Matthias Löwe. Stein's method for dependent random variables occurring in statistical mechanics. *Electron. J. Probab.*, 15:no. 30, 962–988, 2010.
- [Fel68] William Feller. *An introduction to probability theory and its applications. Vol. I*. Third edition. John Wiley & Sons Inc., New York, 1968.
- [GR12] L. Goldstein and G. Reinert. Stein's method for the Beta distribution and the Polya-Eggenberger Urn. [arXiv:1207.1460](#), 2012.
- [LS] C. Ley and Y. Swan. Discrete Stein characterizations and discrete information distances. [arXiv1201.0143](#).
- [Luk94] Ho Ming Luk. *Stein's method for the Gamma distribution and related statistical applications*. ProQuest LLC, Ann Arbor, MI, 1994. Thesis (Ph.D.)—University of Southern California.
- [PR11] Erol A. Peköz and Adrian Röllin. New rates for exponential approximation and the theorems of Rényi and Yaglom. *Ann. Probab.*, 39(2):587–608, 2011.

- [PRR] Erol Peköz, Adrian Röllin, and Nathan Ross. Total variation error bounds for geometric approximation. [arXiv1005.2774](#).
- [SDHR04] Charles Stein, Persi Diaconis, Susan Holmes, and Gesine Reinert. Use of exchangeable pairs in the analysis of simulations. In *Stein's method: expository lectures and applications*, volume 46 of *IMS Lecture Notes Monogr. Ser.*, pages 1–26. Inst. Math. Statist., Beachwood, OH, 2004.
- [Ste72] Charles Stein. A bound for the error in the normal approximation to the distribution of a sum of dependent random variables. In *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability (Univ. California, Berkeley, Calif., 1970/1971), Vol. II: Probability theory*, pages 583–602, Berkeley, Calif., 1972. Univ. California Press.